DOCUMENT RESUME

ED 365 738 TM 021 004

AUTHOR	Bock, R. Darrell; And Others
TITLE	Report on Models for Educational Assessment Involving Multiple-Choice and Free Response Exercises. Open-Ended Exercises in Secondary School Science Assessment. Project 2.4: Quantitative Models To Monitor the Status and Progress of Learning and Performance and Their Antecedents.
INSTITUTION	Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	Nov 93
CONTRACT	R117G10027
NOTE	55p.
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Academic Achievement; Biology; *Educational Assessment; Equivalency Tests; Essay Tests; *Item Analysis; Models; *Multiple Choice Tests; *Science Education; Scoring; Secondary Education; *Test Construction; Test Items; Test Reliability
IDENTIFIERS	Alberta; Alberta Grade Twelve Diploma Examinations; California; *Free Response Test Items; Golden State Examination Program CA; *Open Ended Questions

ABSTRACT

In preparation for a study of essay questions and other forms of open-ended exercises in the California Golden State Examination for biology, the functioning of open-ended biology items in another examination was explored. The Golden State Examination program offers honors credit to students who wish to qualify for admission to programs in California's community college and university system. Test materials and data from the January 1989 grade 12 Alberta (Canada) Diploma Examination in biology were used. A quantitative analysis was made of the relative information about the student's proficiency in biology that is conveyed by the multiple-choice and open-ended items in the Alberta test. An item factor analysis carried out on the full sample of 8,113 examinees revealed the test to be highly unidimensional. A random sample of 2,000 cases was subjected to item response theoretic information analysis. Results suggest that the multiple-choice section has very high reliability and further suggest that one of the open-ended items conveys, on the average, about the same information as four multiple-choice items. The relatively greater efficiency of the multiple-choice items does not, however, imply that they should be used exclusively in measuring science achievement. The implications of the Alberta test for the construction of open-ended questions for the Golden State examination are discussed. Four figures and two tables present results from the test analyses. An appendix contains nine examples of open-ended problems with scoring information. (Contains four references.) (SLD)

ED 365 738

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

National Center for Research on
Evaluation, Standards, and Student Testing

Final Deliverable – November 1993

Project 2.4: Quantitative Models to Monitor the Status and Progress of Learning and Performance and Their Antecedents

Report on Models for Educational Assessment Involving Multiple-Choice and Free Response Exercises

Open-Ended Exercises in Secondary School Science Assessment

R. Darrell Bock, Project Director

U.S. Department of Education
Office of Educational Research and Improvement
Grant No. R117G10027 CFDA Catalog No. 84.117G

Center for the Study of Evaluation
Graduate School of Education
University of California, Los Angeles
Los Angeles, CA 90024-1522
(310) 206-1532

TM021004

The work reported herein was supported under the Educational Research and Development Center Program cooperative agreement R117G10027 and CFDA catalog number 84.117G as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the position or policies of the Office of Educational Research and Improvement or the U.S. Department of Education.

OPEN-ENDED EXERCISES IN SECONDARY SCHOOL SCIENCE ASSESSMENT[1]

R. Darrell Bock, CRESST/University of Chicago

Michele Zimowski, CRESST/National Opinion Research Center

Megan Martin, California Learning Assessment System

Randall Curren, University of Rochester

We are presently engaged in a study of essay questions and other forms of open-ended exercises in the California Golden State Examination for biology. The Golden State Examination program offers honors credit to students who wish to qualify for admission to programs in California's community college and university system. The examination in biology, which is administered to students at the end of their second year in high school, consists of multiple-choice items and open-ended exercises covering main topics in the state curricular guidelines for this subject.

In preparation for our study, we wanted to evaluate the functioning of open-ended items in some other biology honors examination already in use. We were fortunate in this connection to have access to the test materials and data from the January 1989, Grade 12 Alberta Diploma Examination in biology.

A passing score on an Alberta Diploma Examination counts for half credit toward successful completion of a sequence of three courses in the corresponding subject. The examination for the biology sequence is a two-and-one-half hour test made up of 70 multiple-choice items followed by 7 open-ended exercises. The scoring convention, which is explained in the test booklets, awards one mark for each correct response to a multiple-choice item and a total of 30 marks that are assigned to the open-ended items according to correctness and quality of the student's responses.

---

[1] The research reported here was presented at the CRESST conference "Assessment Questions: Equity Answers," University of California, Los Angeles, September 13-14, 1993.

Figure 1 shows an example of one of the open-ended items from the January 1989 examination. The student writes his or her responses to the three parts of the question in complete sentences in the spaces provided. These exercises can be described as "brief" response items: The student is expected to write more than on a short-answer question, but less than on an essay question.

Figure 2 shows the scoring rubric for this item. The scorer assigns one mark to each section if the response consists of an explanation along the lines shown. We refer to this form of scoring as "mark-point" scoring—the student receives marks for certain predetermined points appearing in his or her response.

Item Factor Analysis

We were interested in analyzing quantitatively the relative information about the student's proficiency in biology that is conveyed by the multiple-choice and the open-ended items in this test. As the first step in this analysis, we established that reporting a single, overall score on this test was quite justified by the internal consistency of the test content. A rigorous item factor analysis by the method of Bock, Gibbons, and Muraki (1988) revealed the test to be highly unidimensional. Score variation was dominated by a single principal factor, and there was no hint of any methods factors distinguishing the multiple-choice and the open-ended exercises. The item factor analysis was carried out in the full sample of 8,113 examinees. For the purposes of the item factor analysis, responses to the 7 open-ended items were dichotomized at the midmark.

IRT Test Information Analysis

We next subjected a random sample of 2,000 of these cases to item response theoretic information analysis using the PARSCALE program of Muraki and Bock (1993) (see also Muraki, 1990). This program performs IRT item analysis on graded scores of multiple-category items as well as on binary scores for multiple-choice items in the presence of guessing. It applies the partial-credit model to the graded items and the three-parameter logistic model to the multiple-choice items. The form of partial-credit model used by the program includes parameters for the item discriminating powers,

FOR DEPARTMENT USE ONLY

Use the following chart to answer question 4.

A Comparison of Inhaled and Exhaled Air			
Type of Air	Gas (% by Volume)		
	Oxygen	Carbon Dioxide	Water Vapor
Inhaled Air	20.80	0.04	1.25
Exhaled Air	15.60	4.00	5.90

4. Using complete sentences, explain the difference between inhaled and exhaled air for each gas in the chart. (3 marks)

a. oxygen _____

_____

_____

_____

b. carbon dioxide _____

_____

_____

_____

c. water vapor _____

_____

_____

_____

Figure 1. An open-ended exercise from the January, 1989, Alberta Diploma Examination in biology.

Using complete sentences, explain the difference between inhaled and exhaled air for each gas in the chart.

a. oxygen

1 mark – Some O_2 diffuses from alveoli into the blood during inhalation and thus decreases the amount exhaled.

OR

– Oxygen is used in cellular respiration.

b. carbon dioxide

1 mark – CO_2 diffuses from the blood into the alveoli and is exhaled, thus increasing the amount exhaled.

OR

– CO_2 is produced during cellular respiration.

c. water vapor

1 mark – Exhaled air has been in contact with moist membranes, thus water in the air has increased.

OR

– Air picks up water from the moist nasal passages and the alveoli.

Figure 2. Scoring guide for the exercise in Figure 1.

parameters associated with the thresholds of the items, and parameters for the spacing of the ordered scoring categories. The program estimates all item parameters by the maximum marginal likelihood method and, using these values, estimates a proficiency scale score for each examinee by the maximum likelihood method.

Associated with the maximum likelihood estimation of the proficiency scores is a quantity, called "Fisher information," that evaluates the

contributions of the items to the measurement precision of the test (see Lord, 1980). The information conveyed by the items varies over the proficiency continuum, but we will discuss the information capacity of the items at sample mean proficiency. This makes the results broadly general, since the proficiencies of a large portion of the examinees lie in the neighborhood of the mean. An important property of the test information is that its reciprocal approximates the measurement error variance of the maximum likelihood estimate of proficiency at that point on the scale. If the standard deviation of the scale scores is set to unity, the compliment of the measurement error variance is the true score variance, which in this case is equal to the test reliability. We make use of these relationships in presenting the results of the information analysis in Table 1.

It is apparent from the results in Table 1 that the multiple-choice section of the test has the very high reliability that we would expect from a well-constructed multiple-choice test of 70 items. In fact this IRT reliability is higher than the corresponding Kuder-Richardson reliability because the latter is only a lower bound on reliability and also because, when the item response

Table 1

Information Analysis of the Alberta Diploma Examination in Biology (N=2000)

Section A: 70 Multiple-choice items (Three-parameter logistic model)	
Sample mean	0.0
Sample S.D.	1.0
Information at sample mean	25.00
Reliability	0.960
Section B: 7 Open-ended items (Generalized partial-credit model)	
Sample mean	0.0
Sample S.D.	1.0
Information at sample mean	9.766
Reliability	0.898

model includes discriminating power parameters, the maximum likelihood estimator of examinee proficiency is more precise than the number-right score.

Notice also that the reliability of scores based on the 7 graded items alone would be high enough to justify reporting a separate score for these exercises so as to provide more detailed information on student performance. This level of reliability is achieved with relatively few exercises because graded items, when scored in a consistent manner, are capable of conveying considerably more information than an equal number of multiple-choice items. To show the relative information capacity per item of the multiple-choice and open-ended exercises, we make use of the fact that the item information is additive—that is, the sum of the item information equals the test information. By dividing the test information by the number of items, we therefore obtain the average-per-item information by which the relative information capacity of the two types of items may be compared. Alternatively, we can divide by the amount of time required to respond to the items to obtain a relative measure of item information capacity per unit of testing time. For the latter, we have assumed that the students would spend 80 of the 150 minutes of testing time on the multiple-choice items and 70 minutes on the open-ended items. The results of these relative information comparisons are shown on Table 2. Notice that these results show that one of the open-ended items conveys on average the same information as about four multiple-choice items. However, in terms of the information per minute, an open-ended item has only about 45% of the efficiency of the multiple-choice items.

The relatively greater efficiency of multiple-choice items when testing time is limited does not necessarily imply, however, that they should be used exclusively in measuring achievement in science. The information analysis tells us only about the capacity of the test to make dependable distinctions between students in terms of their position on the proficiency continuum. Although we may be satisfied that the test is operating reliably at the mean level of the students tested, we may not be satisfied that the mean level is high enough to indicate that the science program is accomplishing its objectives of helping students to understand and to explain scientific concepts accurately and clearly in writing. If the examination were to serve the purposes of evaluating this aspect of biology instruction (and not merely to determine the

Table 2

Information Analysis of the Alberta Diploma Examination in Biology—Relative Information (N=2000)

	Relative Information	
	Per item	Per minute[a]
Multiple-choice	0.357	0.313
Open-ended	1.395	0.140
Ratio	3.906	0.447

[a] Assuming 80 minutes for multiple-choice section, 70 minutes for open-ended section.

order of merit among students), then it would be necessary to include the open-ended items despite their lower per-minute efficiency. Assessment implies such evaluation.

Essay Questions for the Golden State Examination in Biology

In our present study, we are attempting to improve the validity of the GSE open-ended items in biology both as measures of student achievement and as indicators of the success of biology instruction in California's high schools. The exercises in our study differ from those of the Alberta Diploma Examination in that they are essay questions rather than brief-answer questions; they focus on broad concepts central to the biology curriculum and require a response that integrates several content topics. The student is allowed 20 minutes, separately timed, to respond to each exercise.

An example of one exercise typical of the nine included in the study is shown in Figure 3. The exercise is fundamentally about how an energy source (in this case, light) can support a self-sustaining ecosystem. We are attempting to evaluate the students' knowledge and understanding of the flow of energy and exchange of matter within the system. The students should describe the flow energy from the light source to the plants and from the plants to the animals; they should mention the carbon-dioxide/oxygen cycle between animals and plants, and refer explicitly or implicitly to the nitrogen cycle in

Suppose that for a biology term project, you and your laboratory partner want to set up an aquarium that can support life for long periods of time without additional care. If you are successful, the fish, plants, and other organisms could live for months or even years without being fed or cared for.

Explain what conditions you and your partner would have to establish in the aquarium to enable the fish to survive for a long time.

How would you set up these conditions?

Explain how each of the following would contribute to continuing life in the aquarium:

several species of tropical fish

snails

bottom-rooted aquatic plants

algae

nitrogen converting bacteria

Figure 3. An exercise from the NORC School Science Assessment Study.

which the bacteria participate. It is anticipated that different students will address these topics at different levels of competency according to their extent of preparation during the biology course and the degree to which they profited from it. In order to capture these levels in the scoring system, we have extended the mark-point method of scoring to include what we call the "graded" mark-point method. Within each point or topic, we identify four levels of competency of response that the scorer can mark. For the exercise in Figure 3, the graded mark-point scoring rubric is shown in Figure 4. In addition to the relatively objective mark-point categories, the scorer is also asked to make a subjective judgment of the overall quality of the response on the 7-point global scale shown at the bottom of Figure 4.

Form No. ☐ – 0 ☐ Paper No. ☐☐☐☐-☐

Exercise 4: "A self-sustaining acquarium"

1. Understanding of the energy flow of the system.

   0 ☐ No reference to light or light sources.

   1 ☐ Refers directly or indirectly to a light source, but does not mention its relationship to plants.

   2 ☐ Mentions that plants will require a source of light.

   3 ☐ Mentions that plants will require a source of light in order to create food for the fish and snails.

   4 ☐ Same as 3, but with greater detail and clarity of explanation.

2. Understanding of the oxygen-carbon dioxide exchange between plants and animals.

   0 ☐ No reference to oxygen and carbon dioxide.

   1 ☐ Mentions need for oxygen and/or carbon dioxide, but refers to an external source or incorrectly identifies the organisms producing these substances.

   2 ☐ Mentions oxygen and/or carbon dioxide and correctly identifies either the oxygen-producing organisms or the carbon dioxide-producing organisms, but not both.

   3 ☐ Mentions both oxygen and carbon dioxide and correctly identifies two or more of the organisms involved in the exchange.

   4 ☐ Same as 3, but with greater detail and clarity of explanation.

3. Understanding of the conversion of matter in the system, including the consumption of plant products by the animals and the recycling of animal and plant waste by the nitrogen-fixing bacteria.

   0 ☐ No mention of exchanges of matter.

   1 ☐ Mentions either animal consumption of plants or bacteria recycling of waste.

   2 ☐ Same as 1, but mentions both processes.

   3 ☐ Same as 2, but with some mention of maintaining the system in a steady state or similar concept.

   4 ☐ Same as 3, but with greater detail and discussion of the interdependence.

4. Specifically mentions the role of the fish, snails, bottom-rooted aquatic plants, algae, and nitrogen-converting bacteria in the above cycles.

   0 ☐ No response or no correct response to this section of the question.

   1 ☐ Some generally accurate description of the role of one or two of the organisms.

   2 ☐ Some generally accurate description of the role of three or four of the organisms.

   3 ☐ Some essentially correct description of the role of all five organisms.

   4 ☐ Same as 3, but with greater detail and clarity of explanation.

5. Global rating: Understanding energy flow and materials exchange between organisms.

   0 ☐ No relevant response: answer sheet is left blank or the student responds "I don't know", etc.

   1 ☐ A meager response showing little or no understanding of the topic.

   2 ☐ A response to some parts of the question; shows some limited understanding of the topic.

   3 ☐ A response to more than half of the question; shows a moderate understanding of the topic.

   4 ☐ A response to most or all of the question; shows a reasonably good understanding of the topic.

   5 ☐ A response to essentially all parts of the question; shows a solid understanding of the topic.

   6 ☐ A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

Figure 4. The scoring rubric for the exercise in Figure 3.

Our study is evaluating the dimensionality and information capacity of the open-ended items relative to each other and to the multiple-choice items of the biology examination. In addition we are evaluating the scoring rubrics in terms of the agreement between raters working independently from written and/or oral instructions and are comparing it with the agreement of the graded mark-point scoring and the global scoring. If our efforts are successful, we hope to be able to formulate rules for writing and scoring open-ended exercises that will improve their operating characteristics and validity for secondary school assessments in biology and other science areas.

References

Bock, R.D., Gibbons, R.D., & Muraki, E. (1988). Full information item factor analysis. *Applied Psychological Measurement, 12*, 261-280.

Lord, F.M. (1980). *Application of item response theory to practical testing problems.* Hillsdale, NJ: Erlbaum.

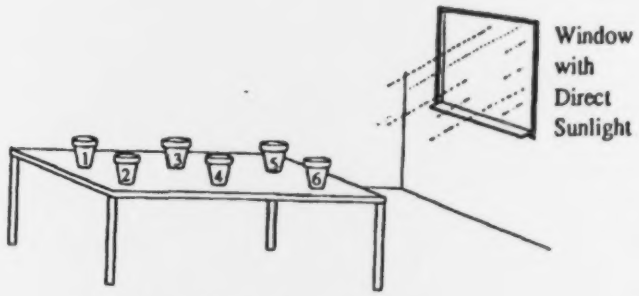
Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement, 14*, 59-71.

Muraki, E., & Bock, R.D. (1993). *PARSCALE: IRT based test scoring and item analysis for graded open-ended exercises and performance tasks.* Chicago, IL: Scientific Software International, Inc.

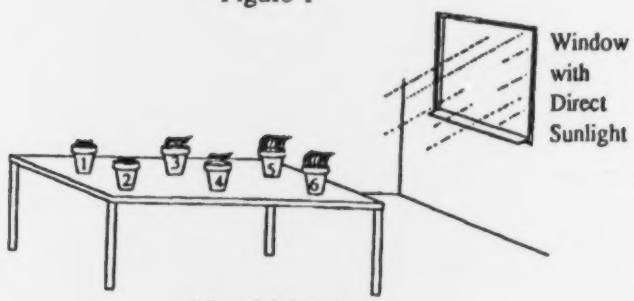
Appendix

Open-ended Problems and Scoring Materials

1. As a biology project, Allison conducted the experiment shown in Figures 1 and 2 below.



At Beginning of Experiment
Figure 1



After 1 Month
Figure 2

She planted grass seeds in 6 small flower pots of the same size using the same type of soil in each pot. She placed the pots on a table along the back wall of the classroom as shown in Figure 1 and watered them as follows:

Pots 1 and 2: 100 mL of H_2O each day

Pots 3 and 4: 100 mL of H_2O every other day

Pots 5 and 6: 100 mL of H_2O once each week

Allison measured the height of the grass in each pot one month after planting. The results of the experiment are shown in Figure 2.

Explain what Allison was testing in her experiment.

Explain what conclusions she could draw from the results.

How could she improve the design of the experiment to make it more valid?

2. Linda is a scientist who wants to develop a form of bacteria that can digest crude oil.

Describe how she might produce and isolate these bacteria.

How would she know whether or not she was successful?

3. For her biology laboratory project, Janet collected bacteria from ten different sources and prepared a microscope slide for each. Her preparations included staining the slides to identify the bacteria. She then examined the slides under a microscope and classified them according to the following criteria:

Criteria	Class of Bacteria	Appearance
1. Shape		
Round	Cocci	
Rods	Bacilli	
Spirals	Spirilla	
2. Clustering		
Single (no clustering)		
Pairing	Diplo-	
Chaining	Strepto-	
Clumping	Staphylo-	
3. Reaction to Staining		
Purple	Positive (+)	
Pink	Negative (−)	

Janet recorded the results of her tests in the following table (a check means she observed the feature).

	Slide	1	2	3	4	5	6	7	8	9	10
Shape:	Round	✓	✓	✓				✓	✓		
	Rods				✓	✓	✓				
	Spirals									✓	✓
Cluster:	Single				✓					✓	✓
	Pairs	✓				✓					
	Chains		✓				✓	✓			
	Clumps			✓					✓		
Stain:	+	✓		✓		✓		✓		✓	
	−		✓		✓		✓		✓		✓

How should Janet have classified the bacteria on slides 2, 5, and 9? From all the observations in the table, which of the following hypotheses are supported? Give reasons for your answers.

1. Spirilli do not cluster.
2. Bacilli always stain positive (+).
3. Cocci and Bacilli do not form clumps.
4. Streptococci never stain positive (+).

4. Suppose that for a biology term project, you and your laboratory partner want to set up an aquarium that can support life for long periods of time without additional care. If you are successful, the fish, plants, and other organisms could live for months or even years without being fed or cared for.

Explain what conditions you and your partner would have to establish in the aquarium to enable the fish to survive for a long time.

How would you set up these conditions?

Explain how each of the following would contribute to continuing life in the aquarium:

several species of tropical fish

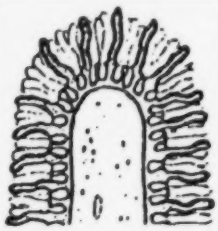
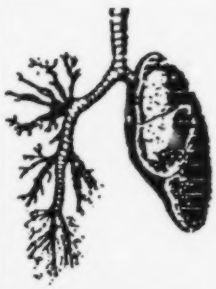
snails

bottom-rooted aquatic plants

algae

nitrogen converting bacteria

5. The following structures are found in plants and animals:

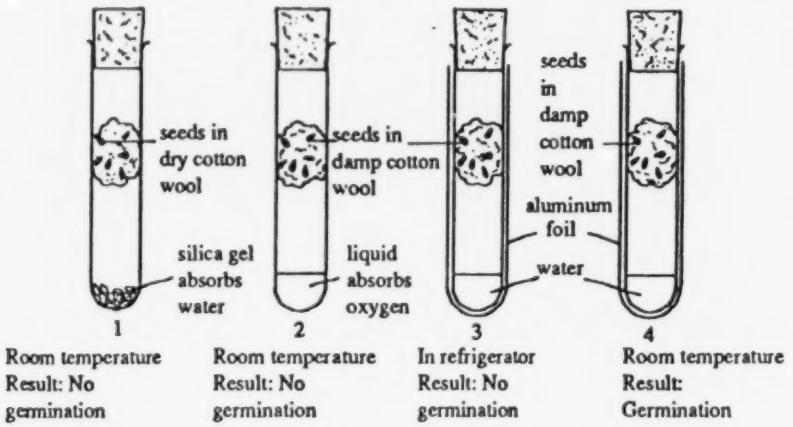
Structure	Name	Where Found
	Root Systems	Plants
	Villi	Small Intestines
	Alveoli	Lungs
	Microtubules	Kidneys

Describe briefly the function of each of the above structures.

Why do these similar structures appear again and again?

Explain the basic principle that is involved in their function.

6. Suppose your classmate Steven set up the following experiment and recorded the results shown below.



Explain what Steven appears to be trying to prove with this experiment.

What conclusions can Steven draw from the results?

How should he improve the experiment to make his conclusions more complete?

7. The following data were obtained in a fifty year study of a woodlot. They show the population density of two tree species, the sugar maple, (*Acer saccharum*), and the hemlock (*Tsuga canadensis*).

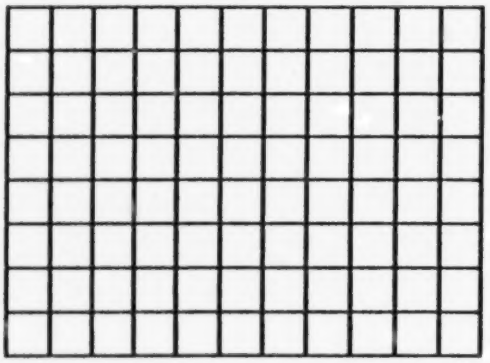
YEAR	1930	1935	1940	1945	1950	1955
A. saccharum	35.0	35.8	36.0	37.4	37.2	37.4
T. canadensis	2.5	4.2	6.3	6.9	7.2	8.8

YEAR	1960	1965	1970	1975	1980
A. saccharum	36.1	34.3	33.1	32.3	30.2
T. canadensis	9.6	10.1	11.9	12.1	12.6

Make a correctly labeled graph of the data on the grid provided below.

Interpret the results, giving two possible hypotheses to account for the trends observed.

What changes in the woodlot over the next 50 years would disprove either of your hypotheses?



8. A biologist observes the following groups of organisms in a small ecosystem:

rabbits	lizards
grasses	fungi
hawks	bacteria
snakes	mice

In the space below, draw a diagram of a food web that includes all of the groups of organisms listed above.

Select one of these groups of organisms in the food web and explain the immediate and long-term effects on the ecosystem if that group of organisms were removed from the food web.

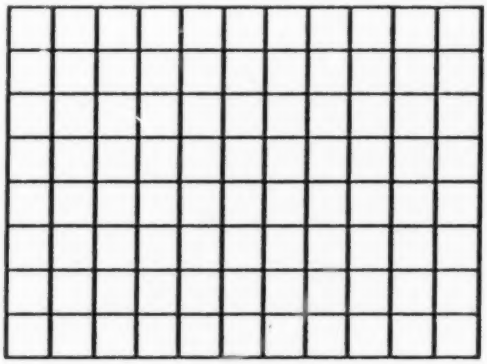
9. In a biological study, 40 grasshoppers were collected in a meadow, and their lengths were measured and recorded. The lengths of the specimens ranged from 11 millimeters (mm) to 90 mm. When these measurements were grouped in intervals of 10 mm, the number of specimens in each interval was as follows:

Length (mm)	Number
11-20	2
21-30	3
31-40	11
41-50	4
51-60	2
61-70	3
71-80	13
81-90	2

Make a correctly labeled graph of the data on the grid provided below.

Carefully examine the graph you prepared. What pattern do you see? Suggest two hypotheses that might account for the pattern of lengths of the grasshoppers.

What could you do to find out whether either of your hypotheses is correct or incorrect?



INSTRUCTIONS FOR RATING RESPONSES TO OPEN-ENDED EXERCISES IN 10TH GRADE BIOLOGY

Thank you for your willingness to participate in the NORC study of open-ended exercises in high school biology. Your assignment is to read and score approximately 200 papers for each of two exercises administered to 10th grade students in May of this year. These students were drawn from biology classes in a random sample of 38 schools in California. The papers you are reading represent a wide range of proficiency in 10th grade biology.

These instructions are intended to be self-contained and suitable for self-study. They include:

- a general description of the rating system,
- copies of the two exercises to which the students were responding,
- copies of the forms on which you are to record your scores, and
- specific instructions for each exercise.

In addition, we have included sample papers (on blue paper) from six students and corresponding score forms indicating what we consider to be good marking of each paper. We believe that these instructions are sufficiently specific and complete to enable you to score the papers according to an objective and consistent standard. If you have any questions about the scoring procedure, please feel to call our "hotline" at 1-800-721-7508. The hotline will be in service from September 1st to September 30th. Someone will be on the line to help you, or an answering machine will take questions that we will answer at a time convenient to you. Please leave your questions along with your telephone number and the best time to reach you.

If you find any serious errors or ambiguities in the scoring guides, please report the problem to the above number so that we can alert the other readers.

General Features of the Rating System

A goal of the rating system is to obtain judgments of the student's understanding of basic principles and concepts in biological science. We want the ratings to be as consistent as possible between independent scorers and from one student to another. The system is designed for use in an assessment program that evaluates the outcome of instruction in the state's public schools and also gives each student a score for his or her performance on the exercises.

To accommodate these uses of the results, the scoring forms have two sections: the first is "semi-analytic" in which subsections of the exercise are rated separately; in the second, a global score is assigned to the student for the overall competence of his or her response to the exercise. Because the students' responses are so varied, it is impossible for the scoring to be completely specified. We therefore depend on your expert judgment of how competently the student has responded to topics within the exercise, or to the exercise as a whole. In some cases, the student's lack of writing skills may obscure his or her actual understanding of the topic. In those cases you will have to make allowances for poor grammar and spelling, and score the understanding of the biology and not the writing skills.

How to Use the Scoring Form

Please look at the sample scoring forms in Appendix 2. The scores for each paper are to be marked on copies of the forms supplied to you. We have enclosed in your package enough copies for all the papers, plus a few extra for practice or in case any are spoiled. At the top of each form are squares in which you must enter the *number of the paper* and the *test form* in which the exercise appeared. These numbers appear at the top right and left, respectively, of each paper. The "A" and "B" following the " – " in the paper number should not be recorded; they refer to the page of the paper. If there is a number following the " – " in the paper number, it should be recorded in the square provided after the " – ". If only an "A" or a "B" follows the "–", the square to the right of the dash on the scoring form should be left blank.

All grades are reported on the form by checking the appropriate boxes. Please use a No. 2 pencil to mark the forms and erase any marks you wish to change. We have left wide margins in case you wish to make any notes about the student's response.

Please return the scoring forms and the papers in two separate piles, but be sure that the order of the forms and the order of the papers are the same in each pile. (See "Returning the Materials" below.) This will help us if we have to refer back to a paper to clarify any marks or comments you may have on the score forms. If possible, return the papers in the same order that they were received.

The Exercises

Please examine the copies of your two exercises shown in Appendix 1. These are exact copies of the exercises presented to the students.

Scoring Section: Semi-analytic

Like the other exercises in this study, those you are scoring have several parts to each question. The upper section of the corresponding scoring sheets contains the rating scales that pertain to each part. Most of these scales have five grades, ranging from no response at all to a very knowledgeable and well-explained response to the topic. The intervening categories represent typical kinds of answers we found in samples of the papers; they describe increasing levels of understanding the topic. Please make the best judgment you can about how the response on a particular paper fits into these categories. To test your judgments, score the six papers in Appendix 3, and refer your ratings to those of our "experts" in Appendix 4.

Scoring Section: Global

After you have completed the analytic scoring of a paper, we ask you to look over the paper again and make a judgment as to what grade it should receive on the seven-category scale in the lower section of the scoring sheet. The definition of the categories is as follows:

- ☐ **0** No relevant response; answer sheet is left blank or student responds "I don't know", etc.
- ☐ **1** A meager response showing little or no understanding of the topic.
- ☐ **2** A response to some parts of the question; shows some limited understanding of the topic.
- ☐ **3** A response to more than half of the question; shows a moderate understanding of the topic.
- ☐ **4** A response to most or all of the question; shows a reasonably good understanding of the topic.
- ☐ **5** A response to all parts of the question; shows a solid understanding of the topic.
- ☐ **6** A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

Apart from the zero category, which corresponds to no response, the remaining six categories are intended to represent more or less equal steps of increasing

understanding of the broad biological principles and concepts that the exercise is designed to evaluate. Specific guidelines for each exercise are found in Appendix 2. Please understand that the present study is purely experimental and the grades that you assign to the paper will not be used for any other purpose. They will not be returned to the school or the student.

Sample Papers

Use some of the extra rating sheets to score the six papers (on blue paper) in Appendix 3. Then compare the results with the ratings in Appendix 4. Always mark the upper semi-analytic section of the rating sheet first and the lower global rating section second. Be sure to record on the answer sheets the form and paper numbers of these sample papers as you read them. Also, write "sample" on these rating sheets.

Time and Comments

Please make a note of the amount of time you spent studying the instructions and the amount of time scoring each of the two exercises. Also, we would appreciate any observations you have about the exercises or the scoring system. We have included a form (on green paper) on which to record this information. We are attempting to develop guidelines for the construction of effective open-ended exercises and a scoring system that will be both reliable and cost-effective. Your careful completion of this phase of the scoring study, and any suggestions you have for improving it, will be a great aid to this effort. Your contribution will be acknowledged by name in the published monograph in which this study will be reported.

Returning the Materials

Once you have finished rating the papers, please pack all of the materials in the box in which you received them. In the envelope marked "Return Postage" you will find a NORC address label and a prepaid postage label. Please affix the address label to the center of the box, the postage label to the upper right-hand corner of the box, and write your return address and "first class" on the box. Please drop it off at a United States Postal Office and inform the attendant that the postage label contains first-class postage for the package.

INSTRUCTIONS FOR RATING RESPONSES TO THE GLOBAL (HOLISTIC) EXERCISES IN 10TH GRADE BIOLOGY

Thank you for your willingness to participate in the NORC study of open-ended exercises in high school biology. Your assignment is to read and score approximately 200 papers for each of two exercises administered to 10th grade students in May of this year. These students were drawn from biology classes in a random sample of 38 schools in California. The papers you are reading represent a wide range of proficiency in 10th grade biology.

These instructions are intended to be self-contained and suitable for self-study. They include:

- a general description of the rating system,
- copies of the two exercises to which the students were responding, and
- specific instructions for each exercise.

In addition, we have included sample papers from six students and corresponding scores indicating what we consider to be good marking of each paper. We believe that these instructions are sufficiently specific and complete to enable you to score the papers according to an objective and consistent standard. If you have any questions about the scoring procedure, please feel to call our "hotline" at 1-800-721-7508. The hotline will be in service from September 1st to September 30th. Someone will be on the line to help you, or an answering machine will take questions that we will answer at a time convenient to you. Please leave your questions along with your telephone number and the best time to reach you.

If you find any serious errors or ambiguities in the scoring guides, please report the problem to the above number so that we can alert the other readers.

General Features of the Global (Holistic) Rating System

A goal of the rating system is to obtain judgments of the student's understanding of basic principles and concepts in biological science. We want

the ratings to be as consistent as possible between independent scorers and from one student to another. The accuracy of the global ratings depends on your expert knowledge of the topics of the exercise and your experience in evaluating students' work in biology. Your task in this rating is to sort the papers for each of the two exercises into six groups representing more or less equal steps of increasing understanding of the broad biological principles and concepts that each exercise is designed to evaluate. To help define these successive groups, we have described them in general terms in the following lists. There is a separate list for each exercise.

EXERCISE 9

Global rating: general understanding of formulating and testing hypotheses based on preliminary data

- ☐ **1** A meager response showing little or no understanding of the topic.
- ☐ **2** A response to some parts of the question; shows some limited understanding of the topic.
- ☐ **3** A response to more than half of the question; shows a moderate understanding of the topic.
- ☐ **4** A response to most or all of the question; shows a reasonably good understanding of the topic.
- ☐ **5** A response to essentially all parts of the question; shows a solid understanding of the topic.
- ☐ **6** A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

EXERCISE 1

Global rating: understanding of principles of good experimental design.

- ☐ **1** A meager response showing little or no understanding of the topic.
- ☐ **2** A response to some parts of the question; shows some limited understanding of the topic.
- ☐ **3** A response to more than half of the question; shows a moderate understanding of the topic.
- ☐ **4** A response to most or all of the question; shows a reasonably good understanding of the topic.
- ☐ **5** A response to essentially all parts of the question; shows a solid understanding of the topic.
- ☐ **6** A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

We have included (on blue paper) six papers that our experts have judged to best represent each of the six levels. After you have read these papers carefully and have an appreciation of the increasing competency they represent, we suggest you use them in the following way to aid your judgments.

Place the six sample papers in order before you on your desk. Then as you read each student paper, place it in a pile behind the sample paper to which it is most similar[1]. When you have sorted in this way as many papers as you can comfortably read in one sitting, look through the piles again briefly to see if you are satisfied with the placements of the papers. When you are satisfied, mark at the top center of each paper (using a No. 2 pencil) the number of the group to which it has been assigned. Then place the piles together in order from 1 to 6 and secure them with a rubber band.

[1]If the student fails to respond to the question or replies with an irrelevant response, such as "I don't know", the paper should be marked with a "0" and set aside in a separate pile.

Proceed with the remaining papers in your next sitting, and with the papers for the second exercise. Please do all of the papers in the first exercise before starting the second exercise. The cover sheets of the exercises tell you which to read first and second.

Time and Comments

Please make a note of the amount of time you spent studying the instructions and the amount of time scoring each of the two exercises. Also, we would appreciate any observations you have about the exercises or the scoring system. We have included a form (on green paper) on which to record this information. We are attempting to develop guidelines for the construction of effective open-ended exercises and a scoring system that will be both reliable and cost-effective. Your careful completion of this phase of the scoring study, and any suggestions you have for improving it, will be a great aid to this effort. Your contribution will be acknowledged by name in the published monograph in which this study will be reported.

Returning the Materials

Once you have finished rating the papers, please pack all of the materials in the box in which you received them. In the envelope marked "Return Postage" you will find a NORC address label and a prepaid postage label. Please affix the address label to the center of the box, the postage label to the upper right-hand corner of the box, and write your return address and "first class" on the box. Please drop it off at a United States Postal Office and inform the attendant that the postage label contains first-class postage for the package.

Form No. ☐ – 0☐ Paper No. ☐☐☐☐☐–☐

Exercise 1: "Allison's grass seed experiment"

1. Explanation of what Allison was testing.

0 ☐ No attempt at explanation.
1 ☐ States only that Allison is studying grass, or how grass grows.
2 ☐ States only that she is studying the effect of amount of sunlight on growth of grass.
3 ☐ States only that she is studying the effect of amount of water or frequency of watering on growth of grass.
4 ☐ Suggests that she could be testing either or both the effect of the amount of water or the effect of the amount of light.
5 ☐ Same as 3 or 4 but with greater detail and clarity of exposition.

2. Recognition that the confounding of the amount of water and the amount of sunlight prevents a valid conclusion.

0 ☐ No mention of what Allison could conclude.
1 ☐ Mentions only that the amount of water has affected the growth of the grass.
2 ☐ States that the grass appears to grow more with less water; does not mention the effect of variation in sunlight.
3 ☐ Explains that the increasing height of the grass could be due either to the different amounts of water or the different intensities of light.
4 ☐ Same as 3, but with greater detail and clarity of explanation.

3. Conditions for an improved experiment.

0 ☐ No plausible response to this part of the question.
1 ☐ States only that Allison should repeat the experiment or do a larger experiment.
2 ☐ States that Allison should control the light or watering in some other way.
3 ☐ Describes specifically how Allison should vary the watering and the sunlight.
4 ☐ Same as 3, but with awareness that both effects could be tested with suitable arrangements of the pots.

Global rating: understanding of principles of good experimental design.

☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc..
☐ 1 A meager response showing little or no understanding of the topic.
☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

BEST COPY AVAILABLE

Form No. ☐ – 0 ☐ Paper No. ☐☐☐☐ – ☐

Exercise 2: "Linda's crude oil eating bacteria"

1. Recognition that evolutionary adaptation in petroleum-rich sites or under laboratory conditions may have already produced traits in bacteria favorable to metabolizing crude oil.

   0 ☐ No mention of this possibility.
   1 ☐ States only that Linda should test many different bacteria for ability to consume crude oil.
   2 ☐ Linda should look in oil fields, etc., for bacteria to test.
   3 ☐ Same as 2, but with some reference to the bacteria being adapted to these sites or ecological niches by natural selection.
   4 ☐ Same as 3, but greater detail and clarity of explanation.

2. Knowledge of the possibility of finding a variant bacterium or inducing a mutation in a population of bacteria that would be favorable to metabolism of crude oil.

   0 ☐ No reference to variation or mutation.
   1 ☐ Refers to finding a favorable variant or mutant, but without explanation.
   2 ☐ Makes some reference to how bacteria populations change through selection or mutation.
   3 ☐ Refers to inducing a favorable mutation by radiation, chemical mutagens, or genetic engineering.
   4 ☐ Same as 2 or 3, but with greater detail and clarity of explanation.

3. Description of methods of identifying, culturing, isolating, and establishing the required strain of bacteria.

   0 ☐ No reference to bacteriological methods.
   1 ☐ Some reference to culturing and culture media.
   2 ☐ Gives some description of how a culture medium is used.
   3 ☐ Same as 2, but includes some explanation of how the oil-eating bacteria are separated from other bacteria.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

4. Proposed test of successfully developing a form of bacteria to digest crude oil.

   0 ☐ No plausible test proposed.
   1 ☐ States only that Linda should check that the bacteria are digesting crude oil.
   2 ☐ Suggests how Linda should carry out tests of the capacity of the bacteria to digest crude oil.
   3 ☐ Same as 2, but includes reference to checking for unfavorable side effects.
   4 ☐ Same as 2 or 3, but with greater detail and clarity of explanation.

Global rating: understanding of sources of biological variation.

☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc.
☐ 1 A meager response showing little or no understanding of the topic.
☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

Form No. ☐ – 0☐ Paper No. ☐☐☐☐☐–☐

Exercise 3: "Janet's classification of bacteria"

1. Classification for slides 2, 5 and 9.

Slide 2. Negative-staining streptococci.

0 ☐ Not mentioned or incorrectly classified.
1 ☐ Classified streptococci, but staining not mentioned or wrong.
2 ☐ Classified streptococci, negative.

Slide 5. Positive staining diplobacilli.

0 ☐ Not mentioned or incorrectly classified.
1 ☐ Classified diplobacilli, but staining not mentioned or wrong.
2 ☐ Classified diplobacilli, positive.

Slide 9. Positive-staining spirilla.

0 ☐ Not mentioned or incorrectly classified.
1 ☐ Classified spirilla, but staining not mentioned or wrong.
2 ☐ Classified spirilla, positive.

2. Supported or not supported hypotheses.
"Spirilla do not cluster"
(Supported: the two spirilla are single.)

0 ☐ Not mentioned.
1 ☐ Incorrect response.
2 ☐ Correct response, no reason given or incorrect reason.
3 ☐ Correct response, correct reason given.

3. "Bacilli always stain positive."
(Not supported: bacilli on slides 4 and 7 are negative-staining.)

0 ☐ Not mentioned.
1 ☐ Incorrect response.
2 ☐ Correct response, no reason given or incorrect reason.
3 ☐ Correct response, correct reason given.

4. "Cocci and bacilli do not form clumps."
(Not supported: the cocci of slide 3 are clumped.)

0 ☐ Not mentioned.
1 ☐ Incorrect response.
2 ☐ Correct response, no reason given or incorrect reason.
3 ☐ Correct response, correct reason given.

5. "Streptococci never stain positive."
(Not supported: the streptococci of slide 7 stain positive.)

0 ☐ Not mentioned.
1 ☐ Incorrect response.
2 ☐ Correct response, no reason given or incorrect reason.
3 ☐ Correct response, correct reason given.

This exercise does not require a global rating

Form No. ☐ – 0☐ Paper No. ☐☐☐☐☐ – ☐

Exercise 4: "A self-sustaining aquarium"

1. Understanding of the energy flow in the system.

   0 ☐ No reference to light or light sources.
   1 ☐ Refers directly or indirectly to a light source, but does not mention its relationship to plants.
   2 ☐ Mentions that plants will require a source of light.
   3 ☐ Mentions that plants will require a source of light in order to create food for the fish and snails.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

2. Understanding of the oxygen-carbon dioxide exchange between plants and animals.

   0 ☐ No reference to oxygen and carbon dioxide.
   1 ☐ Mentions need for oxygen and/or carbon dioxide, but refers to an external source or incorrectly identifies the organisms producing these substances.
   2 ☐ Mentions oxygen and/or carbon dioxide and correctly identifies either the oxygen-producing organisms or the carbon dioxide-producing organisms, but not both.
   3 ☐ Mentions both oxygen and carbon dioxide and correctly identifies two or more of the organisms involved in the exchange.
   4 ☐ Same as 3, but with greater detail and more discussion of the exchanges.

3. Understanding of the conservation of matter in the system, including the consumption of plant products by the animals and the recycling of animal and plant waste by the nitrogen-fixing bacteria.

   0 ☐ No mention of exchanges of matter.
   1 ☐ Mentions either animal consumption of plants or bacteria recycling of waste.
   2 ☐ Same as 1, but mentions both processes.
   3 ☐ Same as 2, but with some mention of maintaining the system in a steady state or similar concept.
   4 ☐ Same as 3, but with greater detail and discussion of the interdependence.

4. Specifically mentions the role of the fish, snails, bottom-rooted aquatic plants, algae, and nitrogen-converting bacteria in the above cycles.

   0 ☐ No response or no correct response to this section of the question.
   1 ☐ Some generally accurate description of the role of one or two of the organisms.
   2 ☐ Some generally accurate description of the role of three or four of the organisms.
   3 ☐ Some essentially correct description of the role of all five organisms.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

Global rating: understanding energy flow and materials exchange between organisms.

☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc.
☐ 1 A meager response showing little or no understanding of the topic.
☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

Form No. ☐ – 0 ☐ Paper No. ☐☐☐☐☐ – ☐

Exercise 5: "Branching structures in organisms"

1. Description of functions of root systems, villi, alveoli, and microtubules.

   0 ☐ No functions described.
   1 ☐ Function of only one of the structures satisfactorily described.
   2 ☐ Functions of two or three of the structures satisfactorily described.
   3 ☐ Functions of all four structures satisfactorily described.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

2. Recognition that the metabolic functioning of all organisms requires transport and exchange of liquids and gasses across membranes.

   0 ☐ Transport, exchange or similar concept not mentioned.
   1 ☐ Concept is mentioned, but only indirectly or inaccurately.
   2 ☐ Transport, exchange or similar common function is mentioned, but not identified as a membrane process.
   3 ☐ Transport, exchange and membranes correctly described and related.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

3. Understanding of the principle that maximizing the ratio of area to volume increases the efficiency of membrane transport.

   0 ☐ No mention of the principle.
   1 ☐ Implied or indirect reference to the importance of a large surface area.
   2 ☐ Clear statement that the structures maximize surface area by branching or folding.
   3 ☐ Specific mention that the maximization involves surface area relative to volume.
   4 ☐ Same as 2 or 3, but with greater detail and clarity of explanation.

Global rating: understanding of transport across membranes and surface area to volume relationships.

☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc.
☐ 1 A meager response showing little or no understanding of the topic.
☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

BEST COPY AVAILABLE

Form No. ☐ – 0 ☐ Paper No. ☐☐☐☐ – ☐

Exercise 6: "Steven's experiment with seeds in cotton wool"

1. Identification of the experimental variables.
(Water, oxygen, light, temperature.)

0 ☐ No reference to any of these variables.
1 ☐ Reference to only one of these variables.
2 ☐ Reference to two or three of these variables.
3 ☐ Reference to all four variables.

2. Explanation of what Steven was trying to prove.

0 ☐ No attempt at explanation.
1 ☐ General statement that Steven was studying conditions affecting seed germination.
2 ☐ Specific mention of the one or two comparisons Steven would make.
3 ☐ Specific mention of three or four of the comparisons Steven would make.
4 ☐ Same as 3, but with greater detail and clarity of explanation.

3. Identification of what the presence or absence of germination in the four test tubes implies about necessary conditions for germination
(Presence of germination only in Tube 4 shows that light is not necessary; comparison of Tubes 3 and 4 shows that room temperature is necessary; other comparisons not conclusive because two conditions are varied simultaneously).

0 ☐ No reference to comparisons between Tube 4 and the other tubes.
1 ☐ Comparisons attempted but the conclusion is incorrect.
2 ☐ Only one of the possible comparisons and conclusions is correctly identified.
3 ☐ The conclusive and inconclusive comparisons in italics above are correctly identified.
4 ☐ Same as 3, but with greater detail and clarity of explanation.

4. Suggestion of treatment combinations that would establish other necessary conditions, assuming that the conditions act independently.

0 ☐ No suggestion of other combinations of conditions.
1 ☐ Statement that other experimental conditions are necessary but they are not identified or are irrelevant.
2 ☐ Suggests relevant additional conditions, but does not identify the appropriate comparisons.
3 ☐ Suggestion that another tube is needed to test whether the difference in germination of Tubes 1 and 4 is due to light or moisture, or between 2 and 4 is due to light or presence of oxygen.
4 ☐ Suggestion of both of these additional conditions, or equivalent comparisons.
5 ☐ Same as 3, but with greater detail and clarity of explanation, possibly including the problem of interactions of two or more conditions.

Global rating: understanding of principles of experimental tests of hypotheses.

☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc.
☐ 1 A meager response showing little or no understanding of the topic.
☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

Form No. ☐ – 0☐ Paper No. ☐☐☐☐☐ – ☐

Exercise 7: "Ecological succession in a wood lot"

1. Graphing the data.

0 ☐ Graph not attempted.
1 ☐ Graph attempted, but one or more data points seriously in error.
2 ☐ Points are correct, but none or only one axis is correctly labeled or the scale interval is poorly chosen.
3 ☐ Points are correct, both axes are correctly labeled, and the choice of scale intervals is satisfactory.
4 ☐ Same as 3, but the graph is drawn and labeled with special accuracy and clarity.

2. Explanation of the cause of the trend in the graph: Hypotheses 1 and 2.

	H1:	H2:	
0	☐	☐	No hypothesis suggested.
1	☐	☐	Suggests only *arbitrary* human intervention in the wood lot or similar inappropriate explanation.
2	☐	☐	Suggests some long-term environmental or economic changes, but does not explain the causal connection with the trend in the relative numbers of the two species of tree.
3	☐	☐	Suggests a plausible hypothesis causally connecting a long-term environmental or economic trend with the trend in the relative numbers of the two species of tree.
4	☐	☐	Same as 3, but with greater detail and clarity of explanation.

3. Suggestion of conditions under which future trends in the relative numbers of the tree populations that would contradict Hypotheses 1 and 2.

	H1:	H2:	
0	☐	☐	No suggestion of how the hypothesis could be contradicted.
1	☐	☐	Suggests changes in conditions, but they bear no plausible relationship to the hypothesis.
2	☐	☐	States only that a reversal of the trend, such that the sugar maple increases and the hemlock declines, would contradict the hypothesis, but does not relate it to changes in conditions.
3	☐	☐	Suggests a future change in the trend of increase and decline caused by a change in environmental or economic conditions.
4	☐	☐	Same as 3, but with greater detail and clarity of explanation.

Global rating: understanding of competition between species in the same ecological niche

☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc.
☐ 1 A meager response showing little or no understanding of the topic.
☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

BEST COPY AVAILABLE

Form No. ☐ – 0☐ Paper No. ☐☐☐☐☐ – ☐

Exercise 8: "Food web"

1. Drawing the food web.

   a. **Relationships.**

   0 ☐ No diagram attempted.
   1 ☐ Diagram is fragmentary or thoroughly confused.
   2 ☐ Diagram has two or three of the groups of organisms correctly related.
   3 ☐ Diagram has four to seven of the groups of organisms correctly related.
   4 ☐ Diagram has all but one of the groups of organisms correctly related.
   5 ☐ Diagram has all eight of the groups of organisms correctly related.

   b. **Direction of relationships.**

   0 ☐ Direction of relationships omitted.
   1 ☐ Direction of relationships indicated.

2. Explanation of the immediate effects of removing a specified group of organisms.

   0 ☐ No explanation of immediate effects.
   1 ☐ Attempts an explanation, but incorrectly identifies the group or groups of organisms that would be most immediately affected.
   2 ☐ Identifies correctly the immediately-affected groups of organisms, but gives no explanation of the ecological relationships between the organisms involved.
   3 ☐ Correctly identifies the immediately-affected organisms and explains their ecological relationship with the group removed.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

3. Explanation of the long-term effects of removing a specified group of organisms.

   0 ☐ No explanation of long-term effects.
   1 ☐ Attempts an explanation, but incorrectly identifies the other group or groups of organisms that would be most affected in the long term.
   2 ☐ Correctly identifies the groups of organisms affected over the long term, but does not give a plausible explanation of the ecological relationships between the affected organisms in the system.
   3 ☐ Correctly identifies the organisms affected over the long term and correctly explains their ecological relationship with the group removed.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

Global rating: understanding of predator-prey relationships

☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc.
☐ 1 A meager response showing little or no understanding of the topic.
☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

Form No. ☐ – 0☐ Paper No. ☐☐☐☐–☐

Exercise 4: "A self-sustaining aquarium"

1. Understanding of the energy flow in the system.

   0 ☐ No reference to light or light sources.
   1 ☐ Refers directly or indirectly to a light source, but does not mention its relationship to plants.
   2 ☐ Mentions that plants will require a source of light.
   3 ☐ Mentions that plants will require a source of light in order to create food for the fish and snails.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

2. Understanding of the oxygen-carbon dioxide exchange between plants and animals.

   0 ☐ No reference to oxygen and carbon dioxide.
   1 ☐ Mentions need for oxygen and/or carbon dioxide, but refers to an external source or incorrectly identifies the organisms producing these substances.
   2 ☐ Mentions oxygen and/or carbon dioxide and correctly identifies either the oxygen-producing organisms or the carbon dioxide-producing organisms, but not both.
   3 ☐ Mentions both oxygen and carbon dioxide and correctly identifies two or more of the organisms involved in the exchange.
   4 ☐ Same as 3, but with greater detail and more discussion of the exchanges.

3. Understanding of the conservation of matter in the system, including the consumption of plant products by the animals and the recycling of animal and plant waste by the nitrogen-fixing bacteria.

   0 ☐ No mention of exchanges of matter.
   1 ☐ Mentions either animal consumption of plants or bacteria recycling of waste.
   2 ☐ Same as 1, but mentions both processes.
   3 ☐ Same as 2, but with some mention of maintaining the system in a steady state or similar concept.
   4 ☐ Same as 3, but with greater detail and discussion of the interdependence.

4. Specifically mentions the role of the fish, snails, bottom-rooted aquatic plants, algae, and nitrogen-converting bacteria in the above cycles.

   0 ☐ No response or no correct response to this section of the question.
   1 ☐ Some generally accurate description of the role of one or two of the organisms.
   2 ☐ Some generally accurate description of the role of three or four of the organisms.
   3 ☐ Some essentially correct description of the role of all five organisms.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

Global rating: understanding energy flow and materials exchange between organisms.

☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc.
☐ 1 A meager response showing little or no understanding of the topic.
☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

BEST COPY AVAILABLE

Form No. ☐ – 0☐ Paper No. ☐☐☐☐–☐

Exercise 5: "Branching structures in organisms"

1. Description of functions of root systems, villi, alveoli, and microtubules.

   0 ☐ No functions described.
   1 ☐ Function of only one of the structures satisfactorily described.
   2 ☐ Functions of two or three of the structures satisfactorily described.
   3 ☐ Functions of all four structures satisfactorily described.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

2. Recognition that the metabolic functioning of all organisms requires transport and exchange of liquids and gasses across membranes.

   0 ☐ Transport, exchange or similar concept not mentioned.
   1 ☐ Concept is mentioned, but only indirectly or inaccurately.
   2 ☐ Transport, exchange or similar common function is mentioned, but not identified as a membrane process.
   3 ☐ Transport, exchange and membranes correctly described and related.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

3. Understanding of the principle that maximizing the ratio of area to volume increases the efficiency of membrane transport.

   0 ☐ No mention of the principle.
   1 ☐ Implied or indirect reference to the importance of a large surface area.
   2 ☐ Clear statement that the structures maximize surface area by branching or folding.
   3 ☐ Specific mention that the maximization involves surface area relative to volume.
   4 ☐ Same as 2 or 3, but with greater detail and clarity of explanation.

Global rating: understanding of transport across membranes and surface area to volume relationships.

☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc.
☐ 1 A meager response showing little or no understanding of the topic.
☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

Form No. ☐ – 0 ☐ Paper No. ☐☐☐☐ – ☐

Exercise 6: "Steven's experiment with seeds in cotton wool"

1. Identification of the experimental variables.
(Water, oxygen, light, temperature.)

0 ☐ No reference to any of these variables.
1 ☐ Reference to only one of these variables:
2 ☐ Reference to two or three of these variables.
3 ☐ Reference to all four variables.

2. Explanation of what Steven was trying to prove.

0 ☐ No attempt at explanation.
1 ☐ General statement that Steven was studying conditions affecting seed germination.
2 ☐ Specific mention of the one or two comparisons Steven would make.
3 ☐ Specific mention of three or four of the comparisons Steven would make.
4 ☐ Same as 3, but with greater detail and clarity of explanation.

3. Identification of what the presence or absence of germination in the four test tubes implies about necessary conditions for germination
(Presence of germination only in Tube 4 shows that light is not necessary; comparison of Tubes 3 and 4 shows that room temperature is necessary; other comparisons not conclusive because two conditions are varied simultaneously).

0 ☐ No reference to comparisons between Tube 4 and the other tubes.
1 ☐ Comparisons attempted but the conclusion is incorrect.
2 ☐ Only one of the possible comparisons and conclusions is correctly identified.
3 ☐ The conclusive and inconclusive comparisons in italics above are correctly identified.
4 ☐ Same as 3, but with greater detail and clarity of explanation.

4. Suggestion of treatment combinations that would establish other necessary conditions, assuming that the conditions act independently.

0 ☐ No suggestion of other combinations of conditions.
1 ☐ Statement that other experimental conditions are necessary but they are not identified or are irrelevant.
2 ☐ Suggests relevant additional conditions, but does not identify the appropriate comparisons.
3 ☐ Suggestion that another tube is needed to test whether the difference in germination of Tubes 1 and 4 is due to light or moisture, or between 2 and 4 is due to light or presence of oxygen.
4 ☐ Suggestion of both of these additional conditions, or equivalent comparisons.
5 ☐ Same as 3, but with greater detail and clarity of explanation, possibly including the problem of interactions of two or more conditions.

Global rating: understanding of principles of experimental tests of hypotheses.

☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc.
☐ 1 A meager response showing little or no understanding of the topic.
☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

Form No. ☐ – 0☐ Paper No. ☐☐☐☐☐ – ☐

Exercise 7: "Ecological succession in a wood lot"

1. Graphing the data.

0 ☐ Graph not attempted.
1 ☐ Graph attempted, but one or more data points seriously in error.
2 ☐ Points are correct, but none or only one axis is correctly labeled or the scale interval is poorly chosen.
3 ☐ Points are correct, both axes are correctly labeled, and the choice of scale intervals is satisfactory.
4 ☐ Same as 3, but the graph is drawn and labeled with special accuracy and clarity.

2. Explanation of the cause of the trend in the graph: Hypotheses 1 and 2.

	H1:	H2:	
0	☐	☐	No hypothesis suggested.
1	☐	☐	Suggests only *arbitrary* human intervention in the wood lot or similar inappropriate explanation.
2	☐	☐	Suggests some long-term environmental or economic changes, but does not explain the causal connection with the trend in the relative numbers of the two species of tree.
3	☐	☐	Suggests a plausible hypothesis causally connecting a long-term environmental or economic trend with the trend in the relative numbers of the two species of tree.
4	☐	☐	Same as 3, but with greater detail and clarity of explanation.

3. Suggestion of conditions under which future trends in the relative numbers of the tree populations that would contradict Hypotheses 1 and 2.

	H1:	H2:	
0	☐	☐	No suggestion of how the hypothesis could be contradicted.
1	☐	☐	Suggests changes in conditions, but they bear no plausible relationship to the hypothesis.
2	☐	☐	States only that a reversal of the trend, such that the sugar maple increases and the hemlock declines, would contradict the hypothesis, but does not relate it to changes in conditions.
3	☐	☐	Suggests a future change in the trend of increase and decline caused by a change in environmental or economic conditions.
4	☐	☐	Same as 3, but with greater detail and clarity of explanation.

Global rating: understanding of competition between species in the same ecological niche

☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc.
☐ 1 A meager response showing little or no understanding of the topic.
☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

Form No. ☐ – 0 ☐ Paper No. ☐☐☐☐☐ – ☐

Exercise 8: "Food web"

1. Drawing the food web.

   a. **Relationships.**

   0 ☐ No diagram attempted.
   1 ☐ Diagram is fragmentary or thoroughly confused.
   2 ☐ Diagram has two or three of the groups of organisms correctly related.
   3 ☐ Diagram has four to seven of the groups of organisms correctly related.
   4 ☐ Diagram has all but one of the groups of organisms correctly related.
   5 ☐ Diagram has all eight of the groups of organisms correctly related.

   b. **Direction of relationships.**

   0 ☐ Direction of relationships omitted.
   1 ☐ Direction of relationships indicated.

2. Explanation of the immediate effects of removing a specified group of organisms.

   0 ☐ No explanation of immediate effects.
   1 ☐ Attempts an explanation, but incorrectly identifies the group or groups of organisms that would be most immediately affected.
   2 ☐ Identifies correctly the immediately-affected groups of organisms, but gives no explanation of the ecological relationships between the organisms involved.
   3 ☐ Correctly identifies the immediately-affected organisms and explains their ecological relationship with the group removed.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

3. Explanation of the long-term effects of removing a specified group of organisms.

   0 ☐ No explanation of long-term effects.
   1 ☐ Attempts an explanation, but incorrectly identifies the other group or groups of organisms that would be most affected in the long term.
   2 ☐ Correctly identifies the groups of organisms affected over the long term, but does not give a plausible explanation of the ecological relationships between the affected organisms in the system.
   3 ☐ Correctly identifies the organisms affected over the long term and correctly explains their ecological relationship with the group removed.
   4 ☐ Same as 3, but with greater detail and clarity of explanation.

Global rating: understanding of predator-prey relationships

☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc.
☐ 1 A meager response showing little or no understanding of the topic.
☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

BEST COPY AVAILABLE

Form No. ☐ – 0 ☐ Paper No. ☐☐☐☐☐ – ☐

Exercise 9: "Lengths of grasshoppers"

1. Graphing the data.

- 0 ☐ Graph not attempted.
- 1 ☐ Graph attempted, but one or more data points seriously in error.
- 2 ☐ Points are correct, but none or only one axis is correctly labeled or the scale interval is poorly chosen.
- 3 ☐ Points are correct, both axes are correctly labeled, and the choice of scale intervals is satisfactory.
- 4 ☐ Same as 3, but the graph is drawn and labeled with special accuracy and clarity.

2. Recognition and description of the bimodal pattern in the graph.

- 0 ☐ No description of the pattern in the graph.
- 1 ☐ A description is attempted, but the presence of the two modes is not mentioned.
- 2 ☐ States that there are large and small groups of grasshoppers, but does not give the positions of the modes.
- 3 ☐ Indicates that there are two modes in the distribution located at lengths 31–40 mm and 71–80 mm.
- 4 ☐ Same as 3, but with additional relevant description of the pattern.

3. Explanation of the distribution of grasshopper lengths: Hypotheses 1 and 2.

	H1:	H2:	
0	☐	☐	No hypothesis proposed.
1	☐	☐	Attempted explanation, but the hypothesis is implausible.
2	☐	☐	Proposed hypothesis is plausible, but overly general and vague.
3	☐	☐	Proposed hypothesis is plausible and specifically related to factors such as sex, maturation, speciation, or dominant and recessive phenotypes.
4	☐	☐	Same as 3, but with greater detail and clarity of explanation.

4. Testing of Hypotheses 1 and 2.

	H1:	H2:	
0	☐	☐	No test proposed.
1	☐	☐	Suggests only further study or collection of more data.
2	☐	☐	Suggestion of further investigation is plausibly related to the hypothesis, but not an effective test.
3	☐	☐	Suggested investigation is plausibly related to the hypothesis and would provide a decisive test.
4	☐	☐	Same as 3, but with greater detail and clarity of explanation.

Global rating: general understanding of formulating and testing hypotheses based on preliminary data.

- ☐ 0 No relevant response; answer sheet is left blank or student responds "I don't know", etc.
- ☐ 1 A meager response showing little or no understanding of the topic.
- ☐ 2 A response to some parts of the question; shows some limited understanding of the topic.
- ☐ 3 A response to more than half of the question; shows a moderate understanding of the topic.
- ☐ 4 A response to most or all of the question; shows a reasonably good understanding of the topic.
- ☐ 5 A response to essentially all parts of the question; shows a solid understanding of the topic.
- ☐ 6 A complete and thorough response to the question; shows an exemplary understanding of the topic; exposition is clear and knowledgeable.

Exercise 1: "Allison's grass seed experiment"

This exercise is about principles of experimentation; it has a minimum of biological content. The student only needs to know that the amount of water and amount of sunlight will affect the rate at which grass grows. The exercise tests the student's understanding of the principle that the ways in which the experimental conditions are varied must not be confounded. The level of one condition must remain fixed while that of the other condition varies. No valid conclusion can be drawn from Allison's experiment because the amount of water and the amount of sunlight received by the plants vary together. Many students perceive the difficulty and suggest that to make the experiment more valid the pots should have been placed in a row at right angle to the window. Then the effect of the watering could be tested without the confounding effect of the sunlight. But still a better response would be one that suggests arranging the pots so that both the effect of water and the effect of sunlight could be tested. That could be done in two experiments, or with the illustrated experimental set up simply by rotating the table by 90 degrees and retaining the same schedule of watering.

Exercise 2: "Linda's crude oil eating bacteria"

This exercise tests the student's knowledge of basic concepts of biological diversity, as they apply to bacteria. The three most common kinds of answers to this question include discussions of 1) adaptation of bacteria to petroleum-containing ecological niches, 2) the existence of mutant bacteria, or an induced mutation, from which a strain of crude oil-eating bacteria could be propagated, 3) genetic engineering by which a gene could be introduced into a strain of bacteria that would produce an enzyme allowing the bacteria to metabolize crude oil.

The student may refer to one or more of these possibilities in his or her response. In addition, some students may suggest how Linda should go about collecting, culturing, isolating and propagating a suitable strain of bacteria.

Exercise 3: "Janet's classification of bacteria"

This exercise is a short-answer problem-solving problem rather than an essay question. It does not require any biological facts apart from those presented in the exercise. Its purpose is to test the student's ability to read a table of data and draw conclusions from it.

Exercise 4: "A self-sustaining aquarium"

This exercise is fundamentally about how an energy source (in this case, light) can support a self-sustaining ecosystem. Some students may answer the question at a very naive "how-to" level, but what we are looking for is some knowledge and understanding of the flow of energy exchange and matter within the system. The main exchanges that the student should mention are the carbon dioxide-oxygen cycle between animals and plants, and the nitrogen cycle in which the bacteria participate. The student is asked to discuss how the plants and animals mentioned participate in these cycles.

Exercise 5: "Branching structures in organisms"

The four pictures presented in this exercise illustrate organismic structures that transport liquids and gasses across membranes. The students are expected to recognize this basic similarity in the functions of the structures. They are also expected to understand that for transport to be as efficient as possible within the space available for the structure, the surface area must be as large as possible. The branching and folding in the illustrated structures produce the required favorable ratio of area to volume.

The students are also asked to describe briefly the function of each of the structures.

Exercise 6: "Steven's experiment with seeds in cotton wool"

In the experiment Steven was apparently testing how seed germination is affected by 1) room temperature vs. refrigeration, 2) light vs. no light, 3) moisture vs. no moisture, and 4) presence of oxygen vs. absence. Of the four combinations of these conditions set up by Steven, only the tube with room temperature, no light, and presence of moisture and oxygen produced germination.

But the conditions in the four tubes are not enough to test all of these variables, even if they act independently. The student should discuss what can be concluded from comparison of the tube where there is germination with each of the other tubes where there is no germination.

Only the comparison between Tube 3 and 4 involves a change of a single variable—temperature. Comparison of Tubes 1 and 4 involves change of two variables, moisture and light. Comparison of Tubes 2 and 4 also involves two variables, oxygen and light. At least two other tubes would have to set up to test separately the effects of water and the presence of oxygen. If there are interactions between the variables, more tests would be necessary. For example, if the effect of absence of oxygen is different under refrigeration than at room temperature, a further two tubes would be necessary to test such an effect.

The exercise is intended to evaluate how far the student can think through the logic of this type of multiple-factor experiment.

Exercise 7: "Ecological succession in a wood lot"

In this exercise we are interested in the student's ability to graph data and describe the trends in the variables.

The student is then required to draw upon biological knowledge, or more general knowledge, to suggest hypothetical explanations of the trends in the data. We do not accept arbitrary human intervention, such as cutting down the sugar maple trees, as a satisfactory explanation. Biological explanations based on changing climate, or changing populations of animals or other plants that affect the growth of the trees is a higher-scoring response. Plausible long-term economic changes, such as decline of demand for maple sugar and neglect of the sugar maples is a satisfactory hypothesis. The hypotheses should involve the effects of "natural" causes operating over a period of years.

As a future changes that might falsify one or other of the hypotheses, we accept any plausible change of conditions that would alter the trend of changes in the relative numbers of the two species of tree over the next 50 years.

Exercise 8: "Food web"

This exercise requires the student to be familiar with the concept of a food web representing predator-prey and foraging relationships between the plants and animals listed. In drawing the web, the student should include all of the organisms on the list and show which are consumed by which other organisms.

The exercise also asks the student to pick one of these groups of organisms and explain the effects on the populations of the other organisms if it were removed from the ecosystem. The student is also asked to distinguish between immediate effects, which in this context refers primarily to effects on organisms one step away in the food web, and long-term effects, which refers to the long-term equilibrium involving all the members of the ecosystem that remain after the selected group has been removed.

Students can respond to this question at many different levels of detail and are rated according to the completeness and accuracy of their response.

Exercise 9: "Lengths of grasshoppers"

In this exercise we are interested in the student's ability to graph data and describe the trends in the variables.

Most of the students will correctly perceive the two modal lengths of grasshoppers and will suggest one or two of the required hypotheses to explain them. We expect mainly one or two of the following four types of hypotheses: 1) that there are two species of grasshopper, one larger and the other smaller, adapted to two different environments in the meadow or to the presence of some predators; 2) that the two sizes correspond to two broods of grasshoppers present in the meadow, an earlier brood now at mature size, and a later brood still juvenile; 3) that the grasshoppers are dimorphic for size, with one sex larger and one sex smaller; 4) that two phenotypes exist in the population, a possibly larger dominant form, and a smaller recessive form. Any of these hypotheses as well as other plausible explanations are acceptable and are rated according to how well they are formulated.

The student is also asked to suggest how the proposed hypotheses could be tested empirically. Some students will merely say that the investigator should repeat the study and collect more data; that is not an acceptable response. Others, with more or less detail and accuracy, will suggest plausible methods of seeking support for the hypotheses. In general terms, acceptable suggestions for each of the above types of hypotheses would include, but not be limited to, the following: *Type* 1. Looking for other anatomical details that would indicate two species of grasshopper, or observing mating patterns for reproductive isolation of the two groups. *Type* 2. Obtaining a sample of the grasshoppers and watching for signs of growth of the smaller insects. *Type* 3. Sexing samples of the grasshoppers anatomically. *Type* 4. Cross-breeding the large and small grasshoppers and looking for segregation of the length phenotype.

The student's two hypotheses and the proposed ways of investigating the hypotheses are rated for their plausibility and clarity of formulation.